



Current state of genome-scale modeling in filamentous fungi

Brandl, Julian; Andersen, Mikael Rørdam

Published in:
Biotechnology Letters

Link to article, DOI:
[10.1007/s10529-015-1782-8](https://doi.org/10.1007/s10529-015-1782-8)

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Brandl, J., & Andersen, M. R. (2015). Current state of genome-scale modeling in filamentous fungi. *Biotechnology Letters*, 37(6), 1131-1139. <https://doi.org/10.1007/s10529-015-1782-8>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Current state of genome-scale modeling in filamentous fungi

Julian Brandl · Mikael R. Andersen

Received: 12 November 2014 / Accepted: 29 January 2015 / Published online: 21 February 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract The group of filamentous fungi contains important species used in industrial biotechnology for acid, antibiotics and enzyme production. Their unique lifestyle turns these organisms into a valuable genetic reservoir of new natural products and biomass degrading enzymes that has not been used to full capacity. One of the major bottlenecks in the development of new strains into viable industrial hosts is the alteration of the metabolism towards optimal production. Genome-scale models promise a reduction in the time needed for metabolic engineering by predicting the most potent targets in silico before testing them in vivo. The increasing availability of high quality models and molecular biological tools for manipulating filamentous fungi renders the model-guided engineering of these fungal factories possible with comprehensive metabolic networks. A typical fungal model contains on average 1138 unique metabolic reactions and 1050 ORFs, making them a vast knowledge-base of fungal metabolism. In the present review we focus on the current state as well as potential future applications of genome-scale models in filamentous fungi.

Keywords Filamentous fungi · Genome-scale models · Metabolic engineering · Metabolism · Systems biology

Introduction

Filamentous fungi have been used for decades in industrial biotechnology exploiting their ability to utilize various sources of nutrients and tolerating adverse growth conditions. For example, tolerance of low pH and the endogenous property of producing citric acid in high amounts have led to the establishment of *Aspergillus niger* as the major source of citric acid production. Furthermore, reflecting the saprobic lifestyle of many filamentous fungi, they harbor a great variety of biomass-degrading enzymes natively produced in high amounts. Additionally, the large diversity of bioactive compounds produced by filamentous fungi is just being recognized as a valuable reservoir of promising new natural compounds. Exploration of the biosynthetic capabilities of these organisms has been facilitated by the availability of genome sequences, thereby enabling the discovery of secondary metabolite clusters being inactive under standard laboratory conditions.

A key requirement for the transition of a new compound into a viable commercial product is the availability of a host producing the compound in sufficiently high amounts. As the organisms are, in general, not evolutionarily optimized to produce a

J. Brandl · M. R. Andersen (✉)
Department of Systems Biology, Technical University of
Denmark, Søltofts Plads 223, 2800 Kongens Lyngby,
Denmark
e-mail: mr@bio.dtu.dk

J. Brandl
e-mail: jubra@bio.dtu.dk

single compound in optimal amounts, process optimizations as well as genetic modifications have to be performed. The rational approach of modifying the metabolism of an organism in order to improve product output constitutes the field of metabolic engineering. Due to the lack of information in the pre-genomic era, the scope of metabolic engineering has been limited to individual pathways not considering inherent interdependencies in the metabolic network. The availability of genome sequence information provides the opportunity to expand the scope of metabolic engineering to the whole metabolism transforming the field towards systems biotechnology. As the process of metabolic engineering represents a bottleneck in the development of many cell factories, accurate predictions by metabolic models could help to reduce the time and costs involved by guiding the efforts towards the most promising set of modifications.

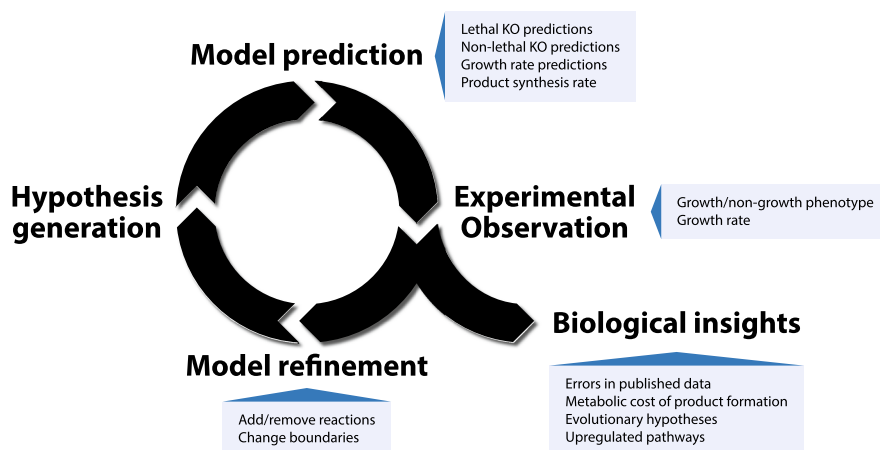
Since the first publication of a complete genome sequence for *Haemophilus influenza* 20 years ago (Fleischmann et al. 1995), the number of published genomes has grown rapidly. The availability of these genomes enabled early reconstruction of the metabolic networks of several species in the groups of viruses (Edwards and Palsson 1999), bacteria (Edwards and Palsson 2000), and yeast (Förster et al. 2003) on the genome scale. These initial drafts have been continuously updated and curated over time, extending the scope and biochemical information contained (Orth et al. 2011; Osterlund et al. 2012). This long period of development demonstrates the iterative nature of model establishment in systems biology where new information is successively included and predictions are validated using experimental data. The process can be visualized as an iterative cycle (see Fig. 1) where the repeated comparison of model predictions with experimental observations leads to biological insights and refinement of the model.

The development of genome-scale network reconstructions (GENRE) in filamentous fungi started considerably later with the first genome being published in 2003 for *Neurospora crassa* (Galagan et al. 2003) followed by the *Aspergillus nidulans* (Galagan et al. 2005), *A. oryzae* (Machida et al. 2005) and *A. fumigatus* (Nierman et al. 2005) in 2005. To date genome-scale reconstructions for the species shown in Table 1 have been published on the basis of available genome sequences and extensive biochemical legacy information.

These reconstructions differ considerably in their content of legacy information included, reflecting different strategies of model establishment. The process of manual reconstruction tends to be laborious, as a maximum amount of information is considered leading to the generation of a structured knowledge base. The complementary approach aims at establishing genome-scale reconstructions (semi-)automatically based on sequence comparisons and gene assignments, enabling the prediction of genome-scale networks for less covered species. While these models generally move towards a larger number of genes and reactions included as models are progressively improved, a qualitative comparison with respect to the numbers of reactions and genes included in the resulting models of these different strategies is not directly possible. The majority of automatically generated models contain dead-ends and/or unconnected reactions that have been removed from manually created and curated models, resulting in higher number of genes included. The results from both of these approaches are fragmentary summaries of the metabolic capacities of the organisms requiring additional curation in the form of gap-filling in order to make computational analysis of the networks feasible. The mathematical background of the underlying modeling approach has been reviewed elsewhere (Llaneras and Picó 2008).

Over the last decade, the conceptual foundations for genome-scale metabolic modelling have been laid by the development of standards for model generation (Le Novère et al. 2005; Thiele and Palsson 2010), model exchange (Hucka et al. 2003) and a variety of computational methods for the analysis of the resulting models (Schellenberger et al. 2011). The labor-intensive nature of model construction led to the development of methods for automating a subset of steps in this process (Agren et al. 2013) up to a fully automated generation of genome-scale metabolic reconstructions (Henry et al. 2010; Pitkänen et al. 2014). This long history of methodical developments now enables researchers to easily generate a genome-scale model for their species of interest and subsequently use it for model-driven discovery and rational strain engineering. The availability of genome sequences as well as of established and well-curated models for many species further facilitates the establishment of new genome-scale reconstructions. As filamentous fungi provide a valuable source of new

Fig. 1 Iterative cycle of model establishment. Examples are taken from Andersen et al. (2009); Melzer et al. (2009); Dreyfuss et al. (2013) and Ledesma-Amaro et al. (2014)



natural products, further increase in the use of genome-scale models in the engineering of these organisms can be expected. In this review we focus on the current state of genome-scale modeling and applications in filamentous fungi and possible future applications.

The applications that have been recorded in filamentous fungi can be divided into three groups. One of the most popular uses for the genome-scale models exploits the underlying collection of information in order to identify gaps in the genome annotation, transfer knowledge to related organisms or generate reduced models for specific analyses. The second category, called model-aided discovery, contains examples where discrepancies between model simulations and experimental data led to the discovery of new biological traits. In the third category, we focus on the applications of genome-scale models for the interpretation of experimental data. This use and the synergism between genome-scale models and omics technologies have been reviewed by Hyduke et al. (2013).

Structural knowledge base of fungal metabolism

Well-curated genome-scale metabolic models represent a structured knowledge base on the metabolism of the particular organism. The link between available genetic, metabolic and bibliomic information represents a valuable resource for researchers that has not been fully exploited. Its potential applications include development of new genome-scale reconstructions,

estimation of network capabilities, as well as analysis of different types of omics data and annotation of related organisms. The enhancement of genome annotations using metabolic networks to represent the relationships between individual enzymatic genes has been coined as “2D annotation” (Reed et al. 2006). The usefulness of this resulting mapping between genes and metabolic functions for the gene function assignment in related organisms has been demonstrated by Liu et al. (2012). The authors used existing genome-scale models as template framework for mapping genes of *Scheffersomyces stipitis*. Similarly, the *Aspergillus niger* genome-scale model (GEM) has been used to assign genes to metabolic pathways in different strains (Andersen et al. 2011).

The existing genome-scale reconstructions for a wide variety of organisms facilitate the development of new metabolic networks as reconstructions of related organisms can be used as a template. This tendency of reusing existing networks is demonstrated by the different network reconstructions of *A. niger*, which have been used as a starting point for the development of the GENRES of *P. chrysogenum* (Agren et al. 2013), *Trichoderma reesei* (Arvas et al. 2011), *Mortierella alpina* and *Mucor circinelloides* (Vongsangnak et al. 2013). Besides this usage for the development of additional genome-scale networks, existing GENRES are actively used as a source to generate reduced models containing all relevant features while facilitating computational analysis. These reduced models can be employed to calculate metabolic fluxes from experimental data as shown by Panagiotou et al. (2008) in their study on the

Table 1 Genome-scale reconstructions of filamentous fungi

Organism	Year	Unique reactions	ORFs	Validation	Reference
<i>Ashbya gossypii</i>	2014	1686 ^b	506	None	Pitkänen et al. (2014)
	2014	1596	766	Growth predictions Riboflavin production	Ledesma-Amaro et al. (2014)
<i>Aspergillus clavatus</i>	2014	2118 ^b	695	None	Pitkänen et al. (2014)
<i>Aspergillus fumigatus</i>	2014	2330 ^b	764	None	Pitkänen et al. (2014)
<i>Aspergillus nidulans</i>	2008	676	666	Growth predictions	David et al. (2008)
	2014	2226 ^b	745	None	Pitkänen et al. (2014)
<i>Aspergillus niger</i>	2008	1190	871	Yield prediction Physiology prediction	Andersen et al. (2008)
	2014	2249 ^b	751	None	Pitkänen et al. (2014)
<i>Aspergillus oryzae</i>	2008	1053	1314	Growth predictions	Vongsangnak et al. (2008)
	2014	2453 ^b	820	None	Pitkänen et al. (2014)
<i>Aspergillus terreus</i>	2013	1357	1454	Growth predictions	Liu et al. (2013)
	2014	2401 ^b	794	None	Pitkänen et al. (2014)
<i>Batrachochytrium dendrobatidis</i>	2014	1979 ^b	556	None	Pitkänen et al. (2014)
<i>Botrytis cinerea</i>	2014	2173 ^b	691	None	Pitkänen et al. (2014)
<i>Chaetomium globosum</i>	2014	1930 ^b	610	None	Pitkänen et al. (2014)
<i>Coprinus cinereus</i>	2014	2080 ^b	636	None	Pitkänen et al. (2014)
<i>Encephalitozoon cuniculi</i>	2014	536 ^b	161	None	Pitkänen et al. (2014)
<i>Fusarium graminearum</i>	2014	2182 ^b	729	None	Pitkänen et al. (2014)
<i>Fusarium oxysporum</i>	2014	2346 ^b	786	None	Pitkänen et al. (2014)
<i>Fusarium verticillioides</i>	2014	2361 ^b	801	None	Pitkänen et al. (2014)
<i>Laccaria bicolor</i>	2014	2162 ^b	666	None	Pitkänen et al. (2014)
<i>Magnaporthe grisea</i>	2014	2152 ^b	686	None	Pitkänen et al. (2014)
<i>Mortierella alpina</i>	2013	1205	1042	None	Vongsangnak et al. (2013)
<i>Mucor circinelloides</i>	2013	1111	1208	None	Vongsangnak et al. (2013)
<i>Mycosphaerella graminicola</i>	2014	2363 ^b	773	None	Pitkänen et al. (2014)
<i>Nectria haematococca</i>	2014	2273 ^b	783	None	Pitkänen et al. (2014)
<i>Neosartorya fischeri</i>	2014	2200 ^b	724	None	Pitkänen et al. (2014)
<i>Neurospora crassa</i>	2013	1027	836	Growth predictions Gene essentiality Nutrient rescue Synthetic lethals	Dreyfuss et al. (2013)
	2014	2189 ^b	691	None	Pitkänen et al. (2014)
<i>Penicillium chrysogenum</i>	2013	1471	1006	None	Agren et al. (2013)
<i>Phaeosphaeria nodorum</i>	2014	2449 ^b	798	None	Pitkänen et al. (2014)
<i>Phanerochaete chrysosporium</i>	2014	2107 ^b	655	None	Pitkänen et al. (2014)
<i>Phycomyces blakesleeanus</i>	2014	2242 ^b	688	None	Pitkänen et al. (2014)
<i>Postia placenta</i>	2014	2047 ^b	647	None	Pitkänen et al. (2014)
<i>Puccinia graminis</i>	2014	1946 ^b	564	None	Pitkänen et al. (2014)
<i>Rhizopus oryzae</i>	2014	2124 ^b	654	None	Pitkänen et al. (2014)
<i>Sclerotinia sclerotiorum</i>	2014	2040 ^b	653	None	Pitkänen et al. (2014)
<i>Trichoderma reesei</i>	2014	2145 ^b	697	None	Pitkänen et al. (2014)
<i>Ustilago maydis</i>	2014	2068 ^b	621	None	Pitkänen et al. (2014)

Table 1 continued

Organism	Year	Unique reactions	ORFs	Validation	Reference
Summary ^a		1912 (1136)	756 1050)		

^a Numbers are averages. Numbers in parenthesis is the average without including the CoReCo-models

^b CoReCo-models; Total number of reactions containing dead-ends

overexpression of phosphoketolase, as well as in the follow-up study on 6-methylsalicylic acid production in *A. nidulans* (Panagiotou et al. 2009).

As genome-scale reconstructions include all theoretically possible metabolic states of an organism, simulating the flux distribution might generate unrealistic results due to the use of otherwise inactive reactions. One example of this is the inability of standard models to accurately predict the secondary metabolite profile produced by a given *Aspergillus* strain. This is partially due to transcriptional and epigenetic regulation of the corresponding gene clusters under standard laboratory conditions (Sanchez et al. 2012). One strategy of dealing with this kind of phenotypic flexibility is the use of omics data as surrogate to determine active reactions in a given environment.

Besides making use of omics data for metabolic modeling, genome-scale reconstructions can be used for the interpretation of omics data by providing the biological framework. There are different examples where genome-scale networks have been used for the analysis of omics data in filamentous fungi. By correlating transcriptional data with calculated flux values during malate production in *A. oryzae* under nitrogen-starvation conditions, Knuf et al. (2013) were able to identify candidates of transcriptionally regulated fluxes representing potential targets for metabolic engineering. Using a reduced model of the *A. niger* GEM, Melzer et al. (2009) developed a similar approach, where they correlated the fluxes through individual pathways with fructofuranosidase production. Using this strategy, they proposed targets for overexpression and gene deletion that would result in an increased fructofuranosidase production.

Model-aided discovery

The potential of GENREs for hypothesis generation starts already at the stage of building the model when a

complete list of reactions present in the organism is compiled. This prediction of missing steps in the metabolic pathways can guide experiments in order to check the predicted function of a gene as well as suggesting functional annotation for orphan genes. This strategy of guiding experimental research using models leads to an iterative cycle refining both the biological knowledge as well as the model itself (see Fig. 1). This use of GEMs for the generation of biological insights is demonstrated by the validation process used in the establishment of the *Neurospora crassa* model (Dreyfuss et al. 2013). Using their reconstruction the authors were able to predict the gene essentially of reactions included in the model with a sensitivity and specificity of 99.1 and 93.6 % respectively. Discrepancy of the predictions with experimental data for the $\Delta erg-14$ mutant led to the discovery of an error in the published knockout strain.

Due to the incomplete characterization of the metabolism for most species, even the most comprehensive of genome-scale reconstructions contain gaps resulting from metabolic reactions not directly characterized in the given organism. However the presence of some specific metabolic reactions can be deduced from observed metabolic capacities of the organism. Identification of these gaps in the network provides the opportunity for guiding the discovery of new genes encoding the enzymatic activity. Candidate genes for individual reactions can be predicted from known homologues in closely related organisms and subsequently tested thereby improving the genome annotation. This kind of strategy has been used by Vongsangnak et al. (2008) to revise the annotation of the *A. oryzae* genome. The authors were able to assign putative functions to 398 newly predicted genes based on gaps identified during the reconstruction of the metabolic network. Similarly, missing reactions in the set-up of the *A. nidulans* GENRE (David et al. 2008) led to the assignment of putative functions to 472 genes based on similarity to orthologous genes catalyzing the missing reaction. This concept of

knowledge-driven gene assignment complements the activity of classical gene function prediction strategies based on sequence motifs. During the generation of genome-scale models, gaps in the set of gene assignments are easily detectable, either by manual curation, examining network connectivity, or modeling (Thiele and Palsson 2010).

Genome-scale model (GEM) aided assessment of phenotypes

Predicting experimental observations not only gives a hint about the outcome of the experiment, but can also give a mechanistic explanation for the observed phenotype. In some of our previous work, we used this approach in the comparison of the citric-acid-production strain *A. niger* ATCC 1015 with the *A. niger* CBS 513.88 used for industrial enzyme production (Andersen et al. 2011). We found an increased expression of genes in the biosynthetic pathways of threonine, serine and tryptophan in the CBS 513.88 strain, reflecting the over-representation of these amino acids in glucoamylase A. Simulation of the fluxes using the GEM of *A. niger* showed that the fluxes carried by these synthetic pathways are required to be at least twice the value in ATCC 1015 in order to support the enzyme production observed in CBS 513.88. This kind of metabolic cost estimation has been proven useful in elucidating the possible reason for the sequence of different acids produced by *A. niger* depending on the ambient pH (Andersen et al. 2009). Considering the acidification potential of the single acids, we have been able to show that the sequence of production is in accordance with an optimal acidification of the surrounding environment at any pH level. This finding suggested that this acidification potential of *A. niger* evolved as a strategy to sustain in a competitive environment.

A similar approach of metabolic cost calculation has been applied to the genome scale model of *A. niger* in order to estimate the metabolic cost associated with by-product formation and reconsumption as shown by (Pedersen et al. 2012). Depending on the magnitude of product formation these kind of metabolic costs are reflected by observable flux changes as demonstrated by Driouch et al. (2012). The authors used the GEM of *A. niger* to determine differences in flux distributions between the wild-type SKANip8 and the fructofuranosidase producing SKAn1015 strain. Using this approach

they showed that expression of the *sucI* gene was sufficient to induce redistribution of fluxes favoring the production of fructofuranosidase. Simulation of the flux capacities predicted the possibility of further increasing fructofuranosidase production up to eight-fold without affecting growth performance.

As data-heavy omics-techniques are becoming increasingly important, the need for tools facilitating the interpretation of the numbers generated from these experiments emerges. Genome-scale models have been used as basis for the construction of organism-specific metabolic maps. Utilizing these maps to plot changes in expression profiles from microarray data provides an intuitive access to the data, as significant changes are visualized and put into the metabolic context. The described approach has been used by Salazar and co-workers in their genome-wide study on glycerol metabolism to analyze changes in the expression of metabolic genes between glycerol and glucose fermentation (Salazar et al. 2009). This strategy allows the immediate visual identification of sub pathways showing increased or decreased activity while comparing different conditions. A similar function is now available in the KEGG database (Kanehisa et al. 2014), where omics data can be loaded and plotted onto the corresponding metabolic network. Tools exploiting the topological information contained in these genome-scale reconstructions have been developed for the identification of highly regulated metabolites in transcriptomic data (Patil and Nielsen 2005).

Discussion

While genome-scale metabolic modeling in *Escherichia coli* and *Saccharomyces cerevisiae* has demonstrated great success in the development and usage of GEMs, the transfer and application to filamentous fungi still lags behind. This delay is not only attributed to the later publication of genome sequences for filamentous fungi, but also reflects the differences in complexity (in particular the number of secreted products and growth morphology) and level of characterization of the individual organisms. A key aspect of metabolism in eukaryotes is the separation of metabolic reactions in different compartments. The correct assignment of metabolic enzymes to these compartments requires detailed information not

readily available for a number of components in filamentous fungi metabolism. Even though this imprecision can limit the accuracy of predictions by these models, they still provide a valuable tool for data interpretation besides being a collection of information that is easily accessible. The unique feature of genome-scale network reconstructions is the integration of different types of information from various sources including databases such as KEGG and Swissprot as well as primary literature data. This ordered collection of metabolic reactions is curated with genomic information and literature citations which results in a structured knowledgebase for the specific organism. While organism specific databases pursue a similar goal, the strength of genome-scale reconstructions is the easy accessibility for computational analysis.

Although genome-scale models have been developed for various filamentous fungi (see Table 1), the directly modeling-driven applications have been limited to date compared to other organisms, in particular considering the large industrial application of these organisms. Possibly due to the complexity of fungi, most published examples focus on applying the models for data interpretation and theoretical calculations. However, these retrospective uses have been relatively successful in providing valuable insights in underlying mechanisms and quantitative assessment of the metabolic state causing a specific phenotype. The potential of these models for metabolic engineering, in contrast, has yet to be demonstrated as most of the published efforts end with the prediction of targets, thereby leaving the proof of feasibility to be made. This lacking verification of model predictions might be attributed to the absence of efficient tools for genetic manipulations. Such tools have long been available in *E. coli* and *S. cerevisiae*, enabling multiple genetic modifications in the same host, but have only recently become available in filamentous fungi (Oakley et al. 2012; Jørgensen et al. 2014; Delmas et al. 2014).

Additional factors potentially hampering the final development of these models into predictive tools for genetic modifications are the low percentage of experimentally verified gene assignments, unknown subcellular localization of the different pathways, as well as insufficient data for model validation. Due to advancements in the field of high throughput omics-technologies this necessary data for modeling is becoming more and more available facilitating the

prospective use of genome-scale models as well as the model building process. The extension of the knowledge being available for incorporation will furthermore improve the prediction accuracy of existing models rendering more advanced applications possible. This improvement of prediction accuracy over time has already been demonstrated by the genome-scale model of *E. coli* where accuracies of >90 and 80 % have been achieved for single- and double gene knockouts respectively (Monk and Palsson 2014). Recent attempts in the prediction of targets for metabolic engineering mainly consisted of finding suitable knock-out and overexpression targets, leaving a gap between the available molecular biological tools and the set of simulated modifications. However it can be expected that with the advent of synthetic biology, the repertoire of modifications will be extended to the introduction of exogenous metabolic reactions and new regulatory circuits further optimizing the network. The increase in accuracy will also provide the possibility to systematically simulate combinations of modifications leading to the assumption free identification of optimal targets for metabolic engineering.

Continuous extension of the models towards more detailed representations while improving their predictive power demonstrates the iterative nature of model development. The availability of predictive models for important organisms used in industrial biotechnology will provide the opportunity to choose the best performing host platform for a new biotechnological process by simulation, as optimal production of an individual compound requires a very specific metabolic performance. This comparability of genome-scale reconstructions between different species has been severely hampered by the non-standardized modeling practice of the different research groups in the past. With the development of modeling standards (Le Novère et al. 2005), a universal data format for model exchange (Hucka et al. 2003) and repositories for existing models (Schellenberger et al. 2010; Henry et al. 2010) as well as naming ambiguities (Bernard et al. 2014), this approach becomes increasingly feasible.

During the past years, life sciences and biotechnology has begun the transformation from data-poor to a data-rich discipline introducing the need for new tools of analysis and interpretation of the generated data. Genome-scale metabolic networks have been proven useful in contextualizing and analyzing the

data acquired by the different omics-techniques. At the same time, the increase in availability and quality of omics data pave the way for the generation of more detailed and context-specific metabolic networks based on the presence of individual metabolic activities. This usage of omics data as a surrogate for the customization of genome-scale models will increase in importance as the sensitivity of the omics technologies is improving. The combination of multi-omics measurements with metabolic modeling provides a powerful tool for the detection and elucidation of unknown relations and interactions. Improvements in the molecular toolbox for genetic engineering will allow for a faster realization of model predictions in the future thereby increasing the speed of model improvement. Taken together the recent developments in the metabolic modeling as well as experimental techniques for generating relevant data are promising for an increased use of model-driven decision making in future.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Agren R, Liu L, Shoaie S et al (2013) The RAVEN toolbox and its use for generating a genome-scale metabolic model for *Penicillium chrysogenum*. PLoS Comput Biol 9:e1002980
- Andersen MR, Nielsen ML, Nielsen J (2008) Metabolic model integration of the bibliome, genome, metabolome and reactome of *Aspergillus niger*. Mol Syst Biol 4:178
- Andersen MR, Lehmann L, Nielsen J (2009) Systemic analysis of the response of *Aspergillus niger* to ambient pH. Genome Biol 10:R47
- Andersen MR, Salazar MP, Schaap PJ et al (2011) Comparative genomics of citric-acid-producing *Aspergillus niger* ATCC 1015 versus enzyme-producing CBS 513.88. Genome Res 21:885–897
- Arvas M, Pakula T, Smit B et al (2011) Correlation of gene expression and protein production rate—a system wide study. BMC Genomics 12:616
- Bernard T, Bridge A, Morgat A et al (2014) Reconciliation of metabolites and biochemical reactions for metabolic networks. Brief Bioinform 15:123–135
- David H, Özçelik IS, Hofmann G, Nielsen J (2008) Analysis of *Aspergillus nidulans* metabolism at the genome-scale. BMC Genom 9:163
- Delmas S, Llanos A, Parrou J-L et al (2014) Development of an unmarked gene deletion system for the filamentous fungi *Aspergillus niger* and *Talaromyces versatilis*. Appl Environ Microbiol 80:3484–3487
- Dreyfuss JM, Zucker JD, Hood HM et al (2013) Reconstruction and validation of a genome-scale metabolic model for the filamentous fungus *Neurospora crassa* using FARM. PLoS Comput Biol 9:e1003126
- Driouch H, Melzer G, Wittmann C (2012) Integration of in vivo and in silico metabolic fluxes for improvement of recombinant protein production. Metab Eng 14:47–58
- Edwards JS, Palsson BO (1999) Systems properties of the *Haemophilus influenzae* Rd metabolic genotype. J Biol Chem 274:17410–17416
- Edwards JS, Palsson BO (2000) The *Escherichia coli* MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. Proc Natl Acad Sci USA 97:5528–5533
- Fleischmann RD, Adams MD, White O et al (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269:496–512
- Förster J, Famili I, Fu P et al (2003) Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. Genome Res 13:244–253
- Galagan JE, Calvo SE, Borkovich KA et al (2003) The genome sequence of the filamentous fungus *Neurospora crassa*. Nature 422:859–868
- Galagan JE, Calvo SE, Cuomo C et al (2005) Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. Nature 438:1105–1115
- Henry CS, DeJongh M, Best AA et al (2010) High-throughput generation, optimization and analysis of genome-scale metabolic models. Nat Biotechnol 28:977–982
- Hucka M, Finney A, Sauro HM et al (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics 19:524–531
- Hyduke DR, Lewis NE, Palsson BØ (2013) Analysis of omics data with genome-scale models of metabolism. Mol Biosyst 9:167–174
- Jørgensen MS, Skovlund DA, Johannesen PF, Mortensen UH (2014) A novel platform for heterologous gene expression in *Trichoderma reesei* (teleomorph *Hypocrea jecorina*). Microb Cell Fact 13:33
- Kanehisa M, Goto S, Sato Y et al (2014) Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acid Res 42:D199–D205
- Knuf C, Nookaew I, Brown SH et al (2013) Investigation of malic acid production in *Aspergillus oryzae* under nitrogen starvation conditions. Appl Environ Microbiol 79:6050–6058
- Le Novère N, Finney A, Hucka M et al (2005) Minimum information requested in the annotation of biochemical models (MIRIAM). Nat Biotechnol 23:1509–1515
- Ledesma-Amaro R, Kerkhoven EJ, Revuelta JL, Nielsen J (2014) Genome scale metabolic modeling of the riboflavin overproducer *Ashbya gossypii*. Biotechnol Bioeng 111:1191–1199
- Liu T, Zou W, Liu L, Chen J (2012) A constraint-based model of *Scheffersomyces stipitis* for improved ethanol production. Biotechnol Biofuels 5:72
- Liu J, Gao Q, Xu N, Liu L (2013) Genome-scale reconstruction and in silico analysis of *Aspergillus terreus* metabolism. Mol Biosyst 9:1939–1948
- Llaneras F, Picó J (2008) Stoichiometric modelling of cell metabolism. J Biosci Bioeng 105:1–11

- Machida M, Asai K, Sano M et al (2005) Genome sequencing and analysis of *Aspergillus oryzae*. *Nature* 438:1157–1161
- Melzer G, Esfandabadi ME, Franco-Lara E, Wittmann C (2009) Flux Design: in silico design of cell factories based on correlation of pathway fluxes to desired properties. *BMC Syst Biol* 3:120
- Monk J, Palsson BO (2014) Genetics. Predicting microbial growth. *Science* 344:1448–1449
- Nierman WC, Pain A, Anderson MJ et al (2005) Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. *Nature* 438:1151–1156
- Oakley CE, Edgerton-Morgan H, Oakley BR (2012) Tools for manipulation of secondary metabolism pathways: rapid promoter replacements and gene deletions in *Aspergillus nidulans*. *Methods Mol Biol* 944:143–161
- Orth JD, Conrad TM, Na J et al (2011) A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism–2011. *Mol Syst Biol* 7:535
- Osterlund T, Nookaew I, Nielsen J (2012) Fifteen years of large scale metabolic modeling of yeast: developments and impacts. *Biotechnol Adv* 30:979–988
- Panagiotou G, Andersen MR, Grotkjaer T et al (2008) Systems analysis unfolds the relationship between the phosphoketolase pathway and growth in *Aspergillus nidulans*. *PLoS One* 3:e3847
- Panagiotou G, Andersen MR, Grotkjaer T et al (2009) Studies of the production of fungal polyketides in *Aspergillus nidulans* by using systems biology tools. *Appl Environ Microbiol* 75:2212–2220
- Patil KR, Nielsen J (2005) Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proc Natl Acad Sci USA* 102:2685–2689
- Pedersen L, Hansen K, Nielsen J et al (2012) Industrial glucoamylase fed-batch benefits from oxygen limitation and high osmolarity. *Biotechnol Bioeng* 109:116–124
- Pitkänen E, Jouhten P, Hou J et al (2014) Comparative genome-scale reconstruction of gapless metabolic networks for present and ancestral species. *PLoS Comput Biol* 10:e1003465
- Reed JL, Famili I, Thiele I, Palsson BO (2006) Towards multidimensional genome annotation. *Nat Rev Genet* 7:130–141
- Salazar M, Vongsangnak W, Panagiotou G et al (2009) Uncovering transcriptional regulation of glycerol metabolism in *Aspergilli* through genome-wide gene expression data analysis. *Mol Genet Genomics* 282:571–586
- Sanchez JF, Somoza AD, Keller NP, Wang CCC (2012) Advances in *Aspergillus* secondary metabolite research in the post-genomic era. *Nat Prod Rep* 29:351–371
- Schellenberger J, Park JO, Conrad TM, Palsson BØ (2010) BiGG: a biochemical genetic and genomic knowledgebase of large scale metabolic reconstructions. *BMC Bioinform* 11:213
- Schellenberger J, Que R, Fleming RMT et al (2011) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat Protoc* 6:1290–1307
- Thiele I, Palsson BØ (2010) A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat Protoc* 5:93–121
- Vongsangnak W, Olsen P, Hansen K et al (2008) Improved annotation through genome-scale metabolic modeling of *Aspergillus oryzae*. *BMC Genomics* 9:245
- Vongsangnak W, Ruenwai R, Tang X et al (2013) Genome-scale analysis of the metabolic networks of oleaginous Zygomycete fungi. *Gene* 521:180–190